Xerox Docket No. D/A2016

ABSTRACT OF THE DISCLOSURE

Systems and methods for triage of passages of text output from an OCR system by use of trainable models of the accuracy of the OCR system based on attributes of individual characters. The systems and methods according to this invention automatically triage an OCR-output text passage by determining at least one OCR-output character attribute for each OCR-output character, determining an error rate for the OCR-output text passage using a triage model and the determined at least one OCR-output character attribute, and comparing the determined error rate for the OCR-output text passage with an OCR-output text passage threshold error rate to perform an OCR-output text passage triage decision. Triage decision includes for example, sending OCR results directly to an end user without any post-OCR processing, sending the OCR results through a post-OCR inspection and processing stage, sending the original document image to be completely keyed in manually, and a combination thereof.